

Comparación de la distribución triangular, con una distribución tipo I de Pearson, como modelos para los problemas en ambiente de riesgo e incertidumbre

Herrerías Velasco, José Manuel jmherrer@ugr.es

Callejón Céspedes, José callejón@ugr.es

Herrerías Pleguezuelo, Rafael rherreri@ugr.es

*Departamento de Métodos Cuantitativos para la Economía y la Empresa
Universidad de Granada*

RESUMEN

Es sobradamente conocido que las distribuciones de probabilidad más utilizadas como modelos probabilísticos, para problemas reales en ambiente de incertidumbre, son la distribución uniforme, la distribución triangular y la distribución beta. Es habitual que la distribución triangular se utilice cuando los tres valores típicos: optimista, pesimista y más probable, suministrados por un experto, pueden considerarse que tienen el mismo interés en el problema, por ello se les da la misma ponderación en la fórmula de la media.

El objetivo principal de este trabajo es seleccionar una distribución del sistema de Pearson, que tenga la misma media que la distribución triangular y comparar ambas distribuciones mediante otras características estocásticas y a través de los diagramas de cociente de momentos, para averiguar cuál de ellas es el modelo probabilístico que mejor se ajusta a los problemas reales en ambiente de riesgo o de incertidumbre.

Palabras claves:

Distribución triangular; sistema de Pearson; ambiente de incertidumbre; diagrama cociente de momentos.

Área temática: A3 Métodos cuantitativos en un entorno con incertidumbre.

XXII Jornadas ASEPUMA – X Encuentro Internacional

Anales de ASEPUMA nº 22: ***Número orden 1301***

ABSTRACT

It is well known that the most commonly probability distributions used as probabilistic models, for real problems in environment of uncertainty, are uniform distribution, triangular distribution and beta distribution. It is usual that the triangular distribution is used when the three typical values: optimistic, pessimistic and most likely values, supplied by an expert, can be considered to have the same interest in the problem, for this reason, are given the same weighting in the formula of the average.

The main objective of this work is to select a distribution of Pearson system, which has the same mean that the triangular distribution and compare both distributions through other stochastic characteristics and moment ratio diagrams, to find out which of them is the probabilistic model that best fits the real problems in environment risk or uncertainty.

Key words:

Triangular distribution; Pearson system; uncertainty; moment ratio diagram.

1. INTRODUCCIÓN

Muchos fenómenos reales tienen una clásica forma de comportarse: empiezan con una etapa de crecimiento, que culmina en un valor máximo, y, a continuación sigue una etapa de decrecimiento. Este comportamiento puede representarse por curvas unimodales y en el caso particular de que las etapas de crecimiento y decrecimiento sean lineales, la representación es un triángulo. También puede considerarse que las ramas de curva creciente y decreciente pueden sustituirse, en una primera aproximación, por funciones lineales. Lo anterior, justifica el pronto interés por el estudio de la distribución triangular, que se remonta al trabajo del profesor de la Real Academia Militar de Wolwich (Inglaterra), Simpson (1755), al que se le deben importantes contribuciones en las Ciencias Actuariales, aunque hay que esperar al nacimiento del método PERT debido a Clark (1962), para que se produzca un verdadero aluvión de trabajos relativos a esta distribución. Entre los más recientes, pueden citarse a Williams (1992), Keefer y Verdini (1993), Johnson (1997) y (2002), etc. Para conocer diferentes trabajos de la distribución triangular, realizados entre los dos hitos anteriormente señalados, puede consultarse a David (1962).

Por otra parte, el sistema de Pearson de distribuciones continuas univariantes, engloba un gran número de distribuciones. En particular, el tipo I contiene a las distribuciones con representación gráfica de campana, es decir, unimodales y con el recorrido definido en un soporte finito, que son dos características determinantes de la distribución triangular. Por ello, se ha seleccionado una distribución de tipo I de Pearson que tenga la misma media que la distribución triangular, para realizar un análisis comparativo entre ambas distribuciones, con el objetivo de dilucidar cuál de ellas es más adecuada, para utilizarla como modelo probabilístico, en los problemas con ambiente de riesgo y de incertidumbre.

El desarrollo del trabajo es el siguiente: en el epígrafe 2 se presenta la distribución tipo I de Pearson estandarizada, esto es, el recorrido de la variable (A, B) se reduce al intervalo $(0, 1)$, junto con sus principales características estocásticas, deducidas de una relación de recurrencia de momentos, de la que se obtiene la expresión general de todos los momentos respecto del origen. El tercer epígrafe se dedica a la

distribución triangular estandarizada y en él, se resumen sus principales características estocásticas, deducidas de unas expresiones generales para el momento respecto al origen de orden n , que resultan ser más prácticas, en la obtención de los momentos respecto a la media, que la utilizada por Kotz y van Dorp (2004).

En el epígrafe 4 se obtienen los coeficientes de variación, asimetría y curtosis, de ambas distribuciones; el punto 5 se dedica a la confección de los dos diagramas de cociente de momentos, el primero que usa los coeficientes de asimetría y curtosis, y el segundo que utiliza los coeficientes de variación y asimetría.

El sexto epígrafe se reserva para los análisis comparativos de las dos distribuciones estudiadas, y, en el epígrafe séptimo se resaltan las conclusiones derivadas de las comparaciones efectuadas, en los puntos anteriores.

2. DISTRIBUCIÓN TIPO I DE PEARSON ESTANDARIZADA

Las distribuciones tipo I de Pearson se caracterizan, en primer lugar, porque el recorrido de la variable está limitado, en segundo lugar, porque generalmente son distribuciones asimétricas y, en tercer lugar, porque la representación gráfica usual, de su función de densidad, tiene forma de campana, es decir, son distribuciones unimodales. Aunque pueden presentar formas de U, J o L, en raras ocasiones. Por ello, son buenas candidatas para modelizar todos aquellos problemas reales que tengan las tres características señaladas, en particular, la duración de una tarea o los flujos netos de caja de una inversión.

Por otra parte, para simplificar las diferentes fórmulas de momentos y diagramas de cociente de ellos, es conveniente utilizar las variables estandarizadas en su recorrido, de forma que, si X es una variable que toma valores en (A, B) , la variable $Z = \frac{X - A}{B - A}$ varía en el intervalo unitario $(0, 1)$.

El sistema de Pearson, particularizado a este caso, Herrerías (1989), será:

$$\frac{f'(z)}{f(z)} = \frac{z - m}{z(z - 1)} \quad (1)$$

donde m es el valor modal estandarizado, es decir, $m = \frac{M - A}{B - A}$

Integrando la ecuación diferencial, de primer orden, lineal y homogénea (1), se obtiene la función de densidad:

$$f(z) = \begin{cases} K z^m (1-z)^{1-m} & 0 < z < 1 \\ 0 & \text{en otro caso} \end{cases} \quad (2)$$

donde la constante normalizadora K, se deduce mediante

$$K = \left[\int_0^1 z^m (1-z)^{1-m} dz \right]^{-1} \quad (3)$$

A partir de (1), se va a obtener la relación de recurrencia de los momentos de la distribución de probabilidad (2), para ello, primero se multiplica por z^n en (1) y, segundo, se integra en (0, 1), obteniéndose:

$$[(n+2)\alpha_{n+1} - (n+1)\alpha_n] = \alpha_{n+1} - m\alpha_n \quad (4)$$

luego, la relación resultante es:

$$\alpha_{n+1} = \frac{1}{n+3} (m+1+n)\alpha_n \quad (5)$$

Puesto que $\alpha_0 = 1$, si en (5) se dan los valores $n = 0, 1, 2$ y 3 , se tienen los cuatro primeros momentos, respecto al origen:

$$\alpha_1 = \frac{1}{3} (m+1)$$

$$\alpha_2 = \frac{1}{4} (m+2)\alpha_1 = \frac{1}{12} (m+2)(m+1) = \frac{1}{12} (m^2 + 3m + 2) \quad (6)$$

$$\alpha_3 = \frac{1}{5} (m+3)\alpha_2 = \frac{1}{60} (m+3)(m+2)(m+1) = \frac{1}{60} (m^3 + 6m^2 + 11m + 6)$$

$$\alpha_4 = \frac{1}{6} (m+4)\alpha_3 = \frac{1}{360} (m+4)(m+3)(m+2)(m+1) = \frac{1}{360} (m^4 + 10m^3 + 35m^2 + 50m + 24)$$

Observando la génesis anterior de los momentos α_n ($n = 1, 2, 3, 4$), se deduce la expresión general para todos ellos

$$\alpha_n = \frac{2}{(n+2)!} \prod_{r=1}^n (m+r), \text{ para } n = 1, 2, \dots \quad (7)$$

Los primeros momentos, respecto a la media, se obtienen a partir de los momentos (6):

$$\mu_2 = \alpha_2 - \alpha_1^2 = \frac{1}{36}(-m^2 + m + 2) = \frac{1}{18} + \frac{m(1-m)}{36} = \frac{2+m(1-m)}{36} = \frac{1}{36}(m+1)(2-m) \quad (8)$$

$$\mu_3 = \alpha_3 - 3\alpha_2\alpha_1 + 2\alpha_1^3 = \frac{1}{270}(2m^3 - 3m^2 - 3m + 2) = \frac{1}{270}(m+1)(m-2)(2m-1) \quad (9)$$

$$\mu_4 = \alpha_4 - 4\alpha_3\alpha_1 + 6\alpha_2\alpha_1^2 - 3\alpha_1^4 = \frac{1}{1080}(m+1)(2-m)(m^2 - m + 4) \quad (10)$$

3. DISTRIBUCIÓN TRIANGULAR ESTANDARIZADA

Se dice que la variable aleatoria Z sigue una distribución triangular $T(0, m, 1)$ si y sólo si su función de densidad es:

$$f(z) = \begin{cases} 2 \frac{z}{m} & \text{si } 0 < z \leq m \\ 2 \frac{1-z}{1-m} & \text{si } m \leq z < 1 \\ 0 & \text{en otro caso} \end{cases} \quad (11)$$

La principal ventaja de esta distribución, cuando se emplea como modelo en problemas en ambiente de incertidumbre, es que queda completamente determinada por los tres típicos valores de la metodología PERT. optimista, más probable y pesimista, cuestión que no se cumple cuando se considera como modelo una distribución Beta, por ello, Johnson (1997) la utiliza como una distribución proxy a la distribución Beta, en Análisis de Riesgo.

Los momentos, respecto al origen, se obtienen fácilmente integrando (11), previamente multiplicada por z^n . La expresión general de ellos, véase Kotz y van Dorp (2004), es:

$$\alpha_n = \frac{2(1-m^{n+1})}{(n+1)(n+2)(1-m)} \quad (12)$$

La expresión (12) es conveniente modificarla, para obtener más fácilmente los sucesivos momentos respecto a la media, en la siguiente

$$\alpha_n = \frac{2}{(n+1)(n+2)} \sum_{r=0}^n m^r = \frac{\sum_{r=0}^n m^r}{\binom{n+2}{2}} \quad (13)$$

Dando valores a $n = 1, 2, 3, 4$, se obtienen los cuatro primeros momentos, respecto al origen:

$$\begin{aligned} \alpha_1 &= \frac{1+m}{3} \\ \alpha_2 &= \frac{1}{6}(1+m+m^2) \\ \alpha_3 &= \frac{1}{10}(1+m+m^2+m^3) \\ \alpha_4 &= \frac{1}{15}(1+m+m^2+m^3+m^4) \end{aligned} \quad (14)$$

Puede comprobarse que la media de la distribución triangular coincide con la media de la distribución de tipo I de Pearson estandarizada; basta observar las expresiones de α_1 obtenidas en (6) y (14).

Los momentos respecto a la media, se obtienen a partir de los momentos dados por (14)

$$\mu_2 = \alpha_2 - \alpha_1^2 = \frac{1}{18}(m^2 - m + 1) = \frac{1}{18}[1 - m(1 - m)] = \frac{1}{18} - \frac{m(1 - m)}{18} \quad (15)$$

$$\mu_3 = \alpha_3 - 3\alpha_2\alpha_1 + 2\alpha_1^3 = \frac{1}{270}(m+1)(m-2)(2m-1) \quad (16)$$

$$\mu_4 = \alpha_4 - 4\alpha_3\alpha_1 + 6\alpha_2\alpha_1^2 - 3\alpha_1^4 = \frac{1}{135}[1 - m(1 - m)]^2 = \frac{1}{135}(m^2 - m + 1)^2 \quad (17)$$

4. COEFICIENTES DE VARIACIÓN, ASIMETRÍA Y CURTOSIS

En este apartado se van a obtener los diferentes coeficientes de variación, asimetría y curtosis de las dos distribuciones consideradas que, en general, dependerán de los valores de m , lo que permitirá la realización del gráfico de los coeficientes de variación y la construcción de los diagramas de cociente de momentos, con la ayuda de la hoja Cálculo Microsoft Excel 2010.

$$CV(P) = \frac{\sqrt{\mu_2}}{\alpha_1} = \frac{1}{2} \sqrt{\frac{2-m}{1+m}} \quad CV(T) = \sqrt{\frac{m^2 - m + 1}{2m^2 + 4m + 2}} \quad (18)$$

$$\beta_1(P) = \frac{\mu_3}{\mu_2^{3/2}} = \frac{4}{5} \frac{1-2m}{\sqrt{(1+m)(2-m)}} \quad \beta_1(T) = \frac{\sqrt{2}}{5} \frac{(2-m)(1+m)(1-2m)}{(m^2 - m + 1)^{3/2}} \quad (19)$$

$$\beta_2(P) = \frac{\mu_4}{\mu_2^2} = \frac{6}{5} \frac{m^2 - m + 4}{(1+m)(2-m)} \quad \beta_2(T) = \frac{12}{5} \quad (20)$$

5. DIAGRAMAS DE COCIENTE DE MOMENTOS

El primer antecedente de uso de los diagramas de cociente (o ratio) de momentos se debe a Craig (1936), aunque su uso generalizado se promueve por Elderton y Johnson (1969), al emplearlo en el estudio de las distribuciones del sistema de Pearson. Este diagrama de cociente de momentos suministra información, sobre el reparto de contribuciones de la asimetría y la curtosis de una familia de distribuciones. Su forma clásica se obtiene representando en el eje de abscisas los valores del coeficiente de asimetría, β_1 , y en el eje de ordenadas, los valores del coeficiente de curtosis, β_2 , definidos en (19) y (20), véase Kotz y Johnson (1985).

Sobre los ejes cartesianos (β_1, β_2) , la región del diagrama situada por debajo de la parábola:

$$\beta_2 = \beta_1^2 + 1 \quad (21)$$

se llama región no factible, ya que en ella no pueden existir puntos del diagrama, debido a que todas las distribuciones de probabilidad verifican, Kotz y Johnson (1985), la relación siguiente:

$$\beta_2 \geq \beta_1^2 + 1 \quad (22)$$

Dados unos datos empíricos, la forma de actuar con el diagrama de cociente de momentos es la siguiente:

- 1º. Se obtiene el par de estimaciones muestrales de β_1 y β_2 .
- 2º. Se sitúa dicho par en el diagrama de cociente de momentos
- 3º. Se considera como familia de distribuciones de probabilidad candidata a modelizar los datos empíricos, aquella que tenga su gráfica más próxima al par de estimaciones.

Para dar una base de comparación consistente, cada familia de distribuciones debe expresarse en términos de los mismos o parecidos parámetros.

Diagramas de cociente de momentos de este tipo, para diferentes familias de distribuciones, incluida la Two Sided Power (TSP), pueden verse en Kotz y van Dorp (2004).

Otro diagrama de cociente de momentos, similar al empleado con las curvas de Pearson, pero que utiliza los coeficientes de variación y de asimetría, fue propuesto por Müller y Vahl (1976). Este diagrama, describe consistentemente la región factible para una familia de distribuciones, en función de los dos coeficientes referenciados. Su representación gráfica, se obtiene situando los valores del coeficiente de variación en abscisas, y los valores del coeficiente de asimetría en ordenadas.

Sin pérdida de generalidad, puede considerarse cada distribución definida en su forma estándar, por uno o dos parámetros de forma. Expresándose estos parámetros en función de los coeficientes seleccionados, se obtiene la relación o relaciones entre ellos. Estas relaciones definen una región factible para cada distribución, en forma de una curva, para las distribuciones de un solo parámetro (gamma, logarítmico normal, etc.) o en forma de área, para la distribuciones de dos parámetros (Beta 1, en sus diferentes formas de campana, jota o U, Beta 2, etc.).

Diagramas de cociente de momentos de este tipo, para diferentes familias de distribuciones, pueden verse en Johnson (2002).

Para una mayor información teórica sobre los dos tipos de cociente de momentos descritos, puede consultarse Vargo et al. (2010), y en Herrerías y Herrerías (2013) se encuentra una aplicación para tres tipos de distribuciones trapezoidales.

6. ANÁLISIS COMPARATIVOS

Es inmediato comprobar que la varianza de la distribución tipo I de Pearson estandarizada es superior en $\frac{1}{12}m(1-m)$, a la varianza de la distribución triangular, para cualquier valor de m , salvo en los valores límites $m = 0$ y $m = 1$, donde ambas son iguales a $\frac{1}{18}$. Para comprobar lo primero, basta restar de la expresión (8), la expresión (15) y para lo segundo basta dar en (8) y (15) los valores límites señalados de m . Véase

el gráfico siguiente, de comparación de varianzas, que corrobora lo afirmado anteriormente.

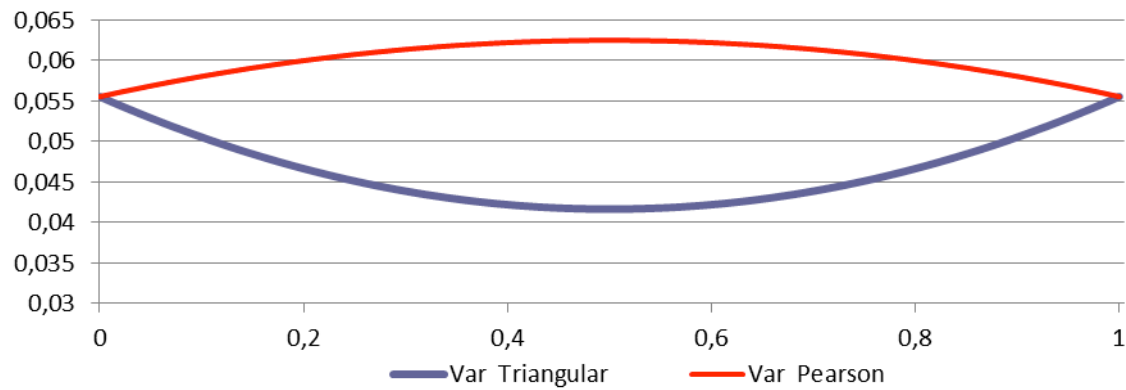


Gráfico 1. Comparación de varianzas

Este resultado trae como consecuencia, que el coeficiente de variación de la distribución tipo I de Pearson, sea siempre superior al de la distribución triangular, salvo en los valores límites $m = 0$ y $m = 1$, donde ambos son coincidentes, véase el siguiente gráfico de comparación de los coeficientes de variación.

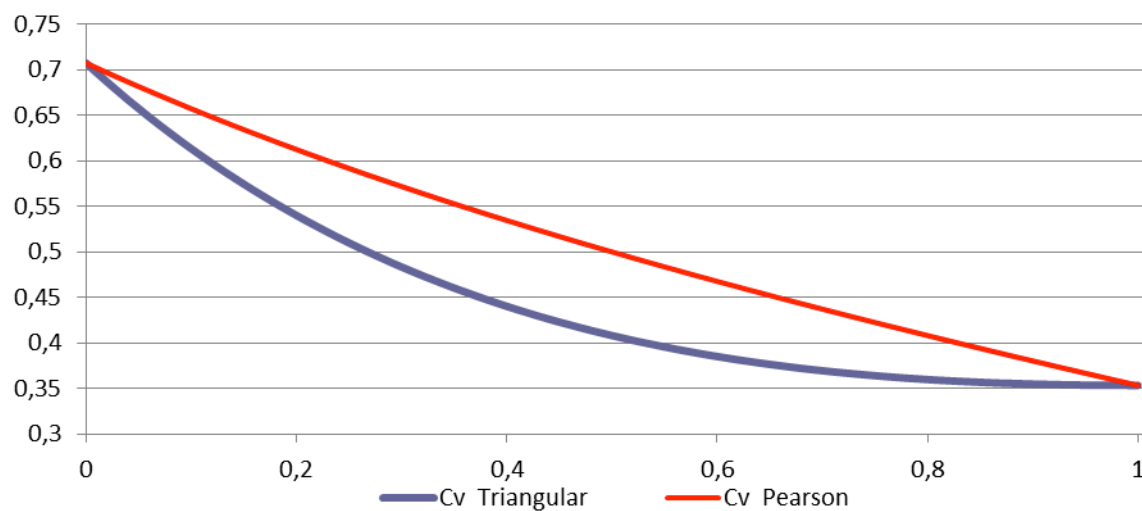


Gráfico 2. Comparación de coeficientes de variación

No deja de ser curioso, que las expresiones de μ_3 dadas en (9) y (16) sean coincidentes, para cualquier valor de m , y como la $\text{Var}(P) \geq \text{Var}(T), \forall m \in (0, 1)$, se tiene como consecuencia, que la asimetría de la distribución triangular será mayor que la asimetría de la distribución tipo I de Pearson.

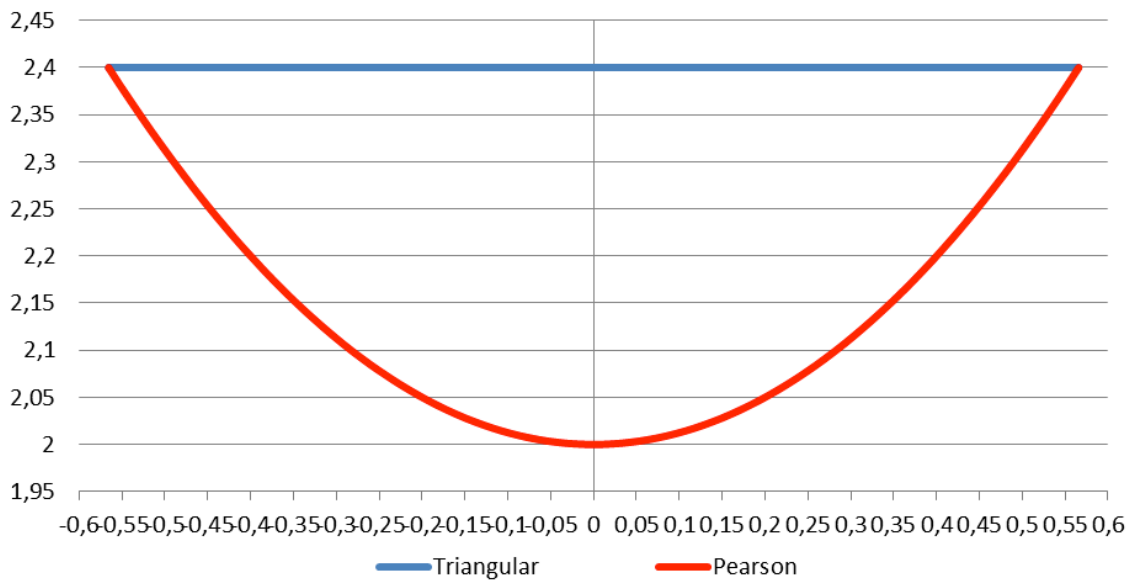


Gráfico 3. Primer diagrama de cociente de momentos

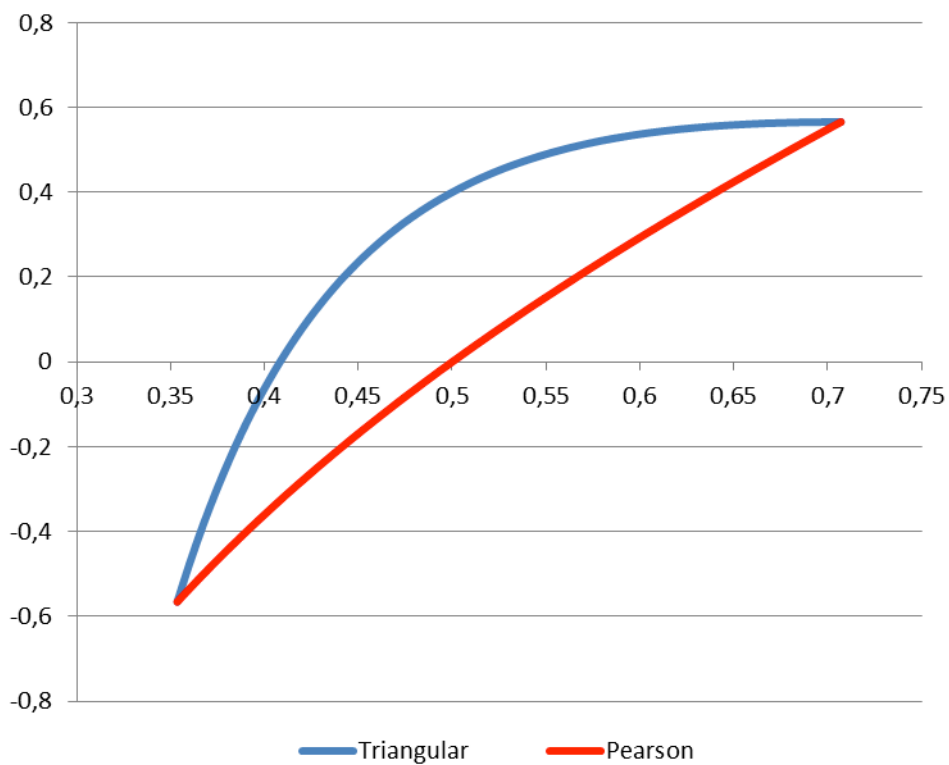


Gráfico 4. Segundo diagrama de cociente de momentos

Al tener la distribución triangular un coeficiente de curtosis constante, su representación gráfica en el diagrama de cociente de momentos (β_1, β_2) es una recta paralela al eje de abscisas y que, por lo señalado en el párrafo anterior, tomará valores superiores a los de la distribución tipo I de Pearson, véase el gráfico del primer diagrama de cociente de momentos.

Por otra parte, como $\beta_2(T) = 12/5 < 3$ y $\beta_2(T) > \beta_2(P)$, por lo anterior, la distribución tipo I de Pearson es más platicúrtica que la distribución triangular.

7. CONCLUSIONES

De acuerdo con los resultados obtenidos, en los puntos anteriores de este trabajo, pueden deducirse las conclusiones parciales siguientes:

1ª) Las distribución tipo I de Pearson y la distribución triangular, quedan perfectamente especificadas con los tres conocidos valores utilizados en el método PERT: (A, M, B), lo que constituye una ventaja sobre la distribución beta $\beta(A, B, p, q)$.

2ª) Como la distribución tipo I de Pearson tiene igual media que la distribución triangular y mayor varianza, es preferible en la modelización de los problemas en ambiente de incertidumbre, consecuentemente con la afirmación de Taha (1981), de que si dos o más rutas tienen la misma media, se elige aquella con la varianza más grande ya que refleja mayor incertidumbre y, por tanto, concluye con resultados más conservadores.

3ª) Si se entiende que la distribución con menor coeficiente de variación es la distribución más adecuada, ya que es la que se comporta más homogéneamente, la distribución triangular es preferible a la distribución tipo I de Pearson.

4ª) Respecto a la flexibilidad para la asimetría, ambas distribuciones son igualmente de flexibles, la misma varía entre (-0,56; 0,56), mientras que la flexibilidad para la curtosis es superior en la distribución tipo I de Pearson, por lo que esta distribución será preferible a la distribución triangular, ya que al ser más flexible, puede modelizar un mayor número de fenómenos reales.

5ª) Fijado el valor del coeficiente de variación, el coeficiente de asimetría de la distribución triangular es superior que el coeficiente de asimetría de la distribución tipo I de Pearson.

6ª) Fijado el valor del coeficiente de asimetría, el coeficiente de variación de la distribución triangular es menor que el coeficiente de variación de la distribución tipo I de Pearson

Como conclusión final se obtiene, que la distribución tipo I de Pearson es más adecuada que la distribución triangular, cuando se trata de modelizar fenómenos reales en ambiente de riesgo o de incertidumbre.

AGRADECIMIENTOS

Este trabajo ha sido desarrollado con la financiación y colaboración del Ministerio de Ciencia e Innovación y la Unión Europea. Proyecto I+D ECO2010-15885; y con ayuda financiera de la Junta de Andalucía, a través del Grupo de Investigación FQM-150: “Modelos Probabilísticos Aplicados a las Ciencias Sociales”.

8. REFERENCIAS BIBLIOGRÁFICAS

- CLARK, C. E. (1962). “The PERT model for the distribution of an activity time”. *Operation Research*. 10,3, pp. 405-406.
- CRAIG, C. G. (1936). “A New Exposition and Chart for the Pearson System of Frequency Curves”. *Annals of Mathematical Statistics*, 7, pp. 16-28.
- DAVID, F. N. (1962). *Games, Gods and Gambling*. Charles Griffin.
- ELDERTON, W. P. y JOHNSON, N. L. (1969). *Systems of Frequency Curves*. Cambridge University Press.
- HERRERÍAS-PLÉGUEZUELO, R. (1989). “Modelos probabilísticos alternativos para el método PERT. Aplicación al Análisis de Inversiones”. *Estudios de Economía Aplicada*. Segunda Reunión Anual de ASEPELT-España. Servicio de Publicaciones de la Universidad de Valladolid, pp. 89-112.

- HERRERÍAS-VELASCO, J. M. y HERRERÍAS-PLÉGUEZUELO, R. (2013). “Comparación de diferentes distribuciones trapezoidales, mediante sus momentos y diagramas de cociente de momentos”. En *Métodos Cuantitativos e Informática. Serie Monografías rect@*, nº4, pp. 159-176.
- JOHNSON, D. (1997). “The triangular distribution as a proxy for the beta distribution in risk analysis”. *The Statistician*, 46, 3, pp. 387-398.
- JOHNSON, D. (2002). “Triangular approximations for continuous random variables in risk analysis”. *Journal of the Operational Research Society*, 53, pp. 457-467.
- KEEFER, D. L. Y VERDINI, A. V. (1993). “Better estimation of PERT activity time parameters”. *Management Science*, 39, 9, pp. 1086-1091.
- KOTZ, S. Y JOHNSON, N. L. (1985). “Moment ratio diagrams”. *Encyclopedia of Statistical Sciences*. Wiley. 5, pp. 602-604.
- KOTZ, S. Y VAN DORP, J. R. (2004). *Beyond Beta. Other Continuous Families of Distributions with Bounded Support and Applications*. World Scientific.
- MÜLLER, P. H. Y VAHL, H. (1976). “Pearson’s system of frequency curves whose left boundary and first three moments are known”. *Biometrika*, 63, pp. 191-194.
- SIMPSON, T. (1755). “A letter to the Right Honourable George Earls of Macesfield. President of the Royal Society, on the advantage of taking the mean of a number of observations in practical astronomy”. *Philosophical Transactions*, 49, 1, pp. 82-93.
- TAHA, H. A. (1981). *Investigación de Operaciones, una introducción. Representaciones y Servicios de Ingeniería S. A. México*.
- VARGO, E., PASUPATHY, R. Y LEEMIS, L. (2010). “Moment-Ratio Diagrams for Univariate Distributions”. *Journal of Quality Technology*, 42, 3, pp. 1-11.
- WILLIAMS, T. M. (1992). “Practical use of distributions in network analysis”. *Journal Operations Research Society*, 43, pp. 265-270.